Annual Reviews in Control 47 (2019) 249–265

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Enabling the human in the loop: Linked data and knowledge in industrial cyber-physical systems

Christos Emmanouilidis[a,*], Petros Pistofidis[b], Luka Bertoncelj[a], Vassilis Katsouros[b], Apostolos Fournaris[c], Christos Koulamas[c], Cristobal Ruiz-Carcel[a]

[a] *Cranfield University, UK*
[b] *Institute for Language and Speech Processing, "Athena" Research Center, Greece*
[c] *Industrial Systems Institute, "Athena" Research Center, Greece*

ARTICLE INFO

Article history:
Received 21 July 2018
Revised 18 January 2019
Accepted 11 March 2019
Available online 19 March 2019

Keywords:
Cyber-physical systems
Internet of things
Context information management
Product lifecycle management
Asset lifecycle management
Maintenance
Human in the loop

ABSTRACT

Industrial Cyber-Physical Systems have benefitted substantially from the introduction of a range of technology enablers. These include web-based and semantic computing, ubiquitous sensing, internet of things (IoT) with multi-connectivity, advanced computing architectures and digital platforms, coupled with edge or cloud side data management and analytics, and have contributed to shaping up enhanced or new data value chains in manufacturing. While parts of such data flows are increasingly automated, there is now a greater demand for more effectively integrating, rather than eliminating, human cognitive capabilities in the loop of production related processes. Human integration in Cyber-Physical environments can already be digitally supported in various ways. However, incorporating human skills and tangible knowledge requires approaches and technological solutions that facilitate the engagement of personnel within technical systems in ways that take advantage or amplify their cognitive capabilities to achieve more effective sociotechnical systems. After analysing related research, this paper introduces a novel viewpoint for enabling human in the loop engagement linked to cognitive capabilities and highlighting the role of context information management in industrial systems. Furthermore, it presents examples of technology enablers for placing the human in the loop at selected application cases relevant to production environments. Such placement benefits from the joint management of linked maintenance data and knowledge, expands the power of machine learning for asset awareness with embedded event detection, and facilitates IoT-driven analytics for product lifecycle management.

1. Introduction

While industrial cyber-physical systems (Colombo, Karnouskos, Shi, Yin, & Kaynak, 2016) bring together the physical and digital worlds in manufacturing, the human integration in production environments has only recently began receiving increased attention (Nunes, Zhang, & Silva, 2015). Terms such as "Operator 4.0" are employed to denote the vision of human empowerment with Industry 4.0 technologies (Romero et al., 2016). Within such a vision, the coexistence of human and engineering actors is viewed through the prism of the nature of their interaction in various forms of physical and digital augmentation of human activity. Many concepts and numerous practical implementation examples of supported human action in industrial environments are reported in the literature. However, the actual cognitive contribution of human activities towards the operation of technical systems, although acknowledged to be important in sociotechnical systems, it remains less well explored.

Arguably, the effectiveness of industrial Human In the Loop (HIL) Cyber Physical Systems (CPS) is linked to the ability to capture and act upon the context of such interaction in an enterprise system (El Kadiri et al., 2016; Nunes et al., 2015). The aim of this paper is to introduce an approach for enabling industrial HIL in CPS, as a contributor to successful integration of sociotechnical systems. Starting form an outline of research efforts related to HIL in CPS with an application focus on product and asset lifecycle management, the paper outlines key emerging HIL-CPS concepts and relevant cognitive capabilities, highlights the role of context information management, and offers examples of placing the HIL at selected relevant application cases. Data and knowledge flows in such activities need to be taken into account and methods for the joint management of linked data and knowledge are introduced as

* Corresponding author.
E-mail addresses: christosem@cranfield.ac.uk, christosem@ieee.org (C. Emmanouilidis), pistofid@ceti-athena.innovation.gr (P. Pistofidis), vsk@ilsp.gr (V. Katsouros), fournaris@isi.gr (A. Fournaris), koulamas@isi.gr (C. Koulamas), cruizcarcel@cranfield.ac.uk (C. Ruiz-Carcel).

https://doi.org/10.1016/j.arcontrol.2019.03.004

a key mechanism for establishing shared context, which in turn is a key ingredient for successful HIL-CPS engagement in industrial environments. The aim is not to seek to replace the human factor but to empower it with more effective integration of human cognitive capabilities within technical systems.

The rest of the paper is structured as follows. Section 2 discusses related work on data flows in Product Lifecycle Management (PLM), HIL in CPS, as well as Linked Data, Knowledge, Context Management and Visual Analytics. Section 3 introduces a linkage between HIL-CPS and human cognitive capabilities and the concept of blending sociotechnical systems with advanced capabilities for interaction, supported by visual analytics and a context management architecture that includes the handling of data, metadata information, and knowledge, together with appropriate business logic. Section 4 presents the key concepts implemented to demonstrate HIL on application cases relevant to product and asset lifecycle management. The first one employs HIL for Linking Data and Knowledge Management and for enhancing machine learning capabilities and integrates operating stage technical systems (condition monitoring) with design stage knowledge (e.g. Failure Modes, Effects, and Criticality Analysis – FMECA). The second case engages HIL with visual analytics to communicate condition monitoring outcomes in visually relevant ways for augmenting human capabilities when performing analytics-driven decision tasks. Section 5 concludes with an outline of the main contributions and pointers for further research.

2. Related work

While much emphasis has been placed on increasing levels of automation in production systems, the case for strengthening the role of HIL is growing stronger, as emerging technology enablers empower human operators to become more effectively integrated in production activities (Romero et al., 2016). Such integration makes human actors and their cognitive capabilities more engaged with data, knowledge, and decision process chains in production environments and PLM activities. Therefore research needs to consider the relevance of data flows when integrating HIL in PLM, the technology enablers that support more effective integration of human cognitive capabilities in CPS, the role of linked data and knowledge in supporting context information management and establishing shared context to facilitate the integration of HIL in technical systems, including the empowering impact of visual analytics as an interactive approach to HIL in decision making. These are discussed next.

2.1. Data flows in PLM

The early vision of closed loop product lifecycle management involved creating information loops between different product lifecycle phases, namely Beginning of Life (BoL), Middle of Life (MoL), and End of Life (EoL) activities (Kiritsis, Bufardi, & Xirouchakis, 2003). Consequent research focused on facilitating information exchanges between the different lifecycle activities (Jun, Kiritsis, & Xirouchakis, 2007). Physical data aspects of such exchanges were handled with radio frequency identification (RFID), introducing the concept of product embedded information devices (PEID) (Kiritsis et al., 2008). This has supported the introduction of smart products or assets (Brintrup et al., 2011; McFarlane, Sarma, Chirn, Wong, & Ashton, 2002; Meyer et al., 2009), a key enabler for the joint handling of operations, maintenance, and logistics (VanBelle et al., 2011) but also of monitoring and control functions (Meyer et al., 2009). However, such joint handling required an upgrade in the level of data exchanges well beyond basic product data ID exchanges. This upgrade could be served by further advancements IoT technologies and collaborative digital engineering systems (Kiritsis, 2011).

Among the key challenges in IoT-enabled product and asset lifecycle management is the integration of data, information and knowledge from disparate and heterogeneous sources. This renders the conventional approach to integrating data through a common enterprise data warehouse increasingly problematic in modern big data enterprise environments (Vathy-Fogarassy & Hugyák, 2017). Instead, the emerging data management pattern is that of retrieving relevant data from disparate sources and seeking to integrate them at the point of end consumption. However, the often heterogeneous nature of the data creates further challenges to such approaches and has led to research efforts to establish semantic interoperability of connected products, supported by developments that looked into how semantic (Cassina et al., 2008) and ontology based modelling (Matsokis and Kiritsis, 2010) can enable such product lifecycle data exchanges. This thread of research has led to standardized, and semantically enhanced product lifecycle data exchanges (Framling et al., 2014; Kubler et al., 2015). IoT connectivity nonetheless gives rise to a multi-layered view of data exchanges, which requires a mapping of the IoT information processing layers with product data modelling, from the physical to the application layer (Framling, Kubler, & Buda, 2014). Product data interoperability therefore is relevant across the IoT stack and involves both lower tier data, such as sensor measurements, but also higher level information (Yoo et al., 2016), in order to enable product data integration in proximity to the point of data consumption.

Among the prime interests in closed loop PLM is feeding MoL information back to BoL activities, so as to enable a better understanding of how specific design choices might perform in operation and drive product design enhancements accordingly. The nature of the required data acquisition, transmission, management, and processing varies from case to case and may create significant challenges. For example, raw sensed data transmission may require considerable bandwidth, while radio frequency (RF) operation may be constrained by the nature of the surrounding environment. The business value of embedded intelligence for smart services was recognized even before the dawn of IoT (Kaplan et al., 2005). Technological solutions involved sensor networking protocols, such as those linked to IEEE 802.15.4, and have been adopted in wireless sensor network applications for asset monitoring (Willig, 2008). Embedded processing of locally acquired measurements on sensor nodes enabled asset intelligence beyond identification and tracking (Liyanage, Lee, Emmanouilidis, & Ni, 2009). Such processing enables assets and products to offer a higher level of self-awareness (Katsouros, Koulamas, Fournaris, & Emmanouilidis, 2015), consistent with an agent-based view of intelligent cyber-physical entities (Leitão et al., 2016). Such cyber-physical monitored assets feature a basic cycle of perception, analysis, decision, and (re)action.

Coupling internetworking connectivity with local, distributed and cloud computing, together with semantically enriching product information, has been recognized as key contributor towards connected and intelligent products in enterprise systems (Kiritsis, 2011). Different terms, such as product avatars (Wuest, Hribernik, & Thoben, 2015), shadows (Vermesan & Friess, 2016), or digital twins (Vermesan et al., 2011) have all been employed to describe the cyber version of a physical asset, acting as a smart agent (Leitão, Member, Ma, & Vrba, 2013), or intelligent product in a cyber-physical world of interconnected physical entities (Leitão et al., 2016). Accompanying physical assets with their cyber counterparts (Wuest et al., 2015) enhance the efficiency of modern enterprise information systems, (El Kadiri et al., 2016). Communicating product lifecycle to the application layer of enterprise systems is best achieved with application layer relevant means, for example with visual product representations. For example, MoL relevant product data can be superimposed to BoL product views, such as a 3D product CAD model (Emmanouilidis, Beroncelj, Bevilacqua, Tedeschi, & Ruiz-Carcel, 2018). Thus, product data are made available to users as a visual product design representation, aiding the understanding of maintenance related concepts, such as the occurrence of failure modes. Such an approach can be useful when dealing with real data streams from products as a natural visual analytics extension to closed loop PLM. What is more, it enables the integration of HIL closing PLM information loops via enriched data and knowledge (Wuest et al., 2015). The next section looks into more detail on how this integration is being pursued in industrial CPS.

2.2. Human in the loop and cyber physical systems

Most research efforts on data exchanges in PLM have focused on automated data exchanges, while human-contributed data and knowledge has received far less attention. However, HIL integration requires further development of methods and technological solutions as it is now being recognised to be a major enabler for CPS (Nunes et al., 2015). Recently, human integration in industrial environments is receiving increasing attention, with terms such as "Operator 4.0" used to denote the vision of human empowerment with Industry 4.0 technologies (Romero et al., 2016). However, while the concept of 'people-centric' IoT has been highlighted in a range of application domains and its cognitive contribution in empowering interaction between human and non-human actors is emphasized in recent studies (Feng, Setoodeh, & Haykin, 2017), the nature of such interaction in industrial environments is still primarily confined to intelligent operator support. Indeed, the actual cognitive contribution of human activities towards the operation of technical systems, although acknowledged to be of importance, is still less well explored.

When considering human integration in CPS, significant added value may arise as a result of the interaction between human and non-human actors. Recent research has proposed the integration of HIL of critical systems and processes, with the role of technical systems being to prevent cognitive overload for human operators, rather than replacing or replicating their function. The approach combined a supervisory loop for situational awareness with a dedicated machine learning approach (Gross et al., 2017). HIL also enhances decision making capabilities. Specifically, a human operator receiving a range of automated decision recommendations, needs to identify an appropriate recommended action and apply it to physical assets. This can result in IoT-driven intelligence, where the iterative nature of the HIL interaction, aided by natural interaction interfaces (e.g. natural language-based, visual analytics), as well via semantically enriched abstraction of data through knowledge, greatly enhance human decision capabilities (Ma et al., 2017), but also machine learning tasks in diagnostics (Subramania & Khare, 2011).

Further scenarios of HIL or more broadly Human in the Mesh involvement in industrial environments have been proposed, including interaction with ERP, MES, SCADA, simulation, analytics, data management, as well as lower level shop floor activities delivering flexibility in CPS-enabled manufacturing environments (Fantini et al., 2016). Identifying appropriate ways to make such interaction more effective and better integrate human with non-human actors is still open to research but methods and tools to align this integration with human cognitive capabilities is a natural promising path. The next section discusses options to achieve such integration.

2.3. Linked data, knowledge, context, and visual analytics

The efficiency of human cognitive activities and therefore of HIL in CPS crucially depends on the human actor having a sound understanding of the context of the targeted problem or situation. In connected factories, human actors, as well as IoT-enabled production equipment and environments create production and product – related streams of data. In order to efficiently manage such data produced by multiple production sites and stakeholders, solutions for scalable data processing are needed. Context Information Management has emerged as a key concept in managing such complexity in IoT-enabled environments (Perera, Zaslavsky, Christen, & Georgakopoulos, 2014). The main principle is that in order to enable efficient aggregation and processing of data from disparate sources, only contextually relevant data need be made available at the point of data or services consumption. While domain specific context greatly varies depending on the targeted application, higher level context can be categorized to fall under certain broad categories, such as asset, user, business, environment and system context.

In industrial CPS, context information management can determine the situational circumstances of decisions (El Kadiri et al., 2016). In product and asset lifecycle management, high-level context can be classified according to the aforementioned broad categories with domain-specific semantics, as illustrated in Fig. 1. Each context category comprises parameters which can be acquired or computed and their semantic interpretation would impact on the way a specific situation needs to be assessed. For example, relevant information and services depend on the asset under consideration. Therein, the *asset context* may be determined by considering the status of the modelled asset in the asset hierarchy, its function within the production system, historical data about its operation, including prognostics and health management (PHM), as well as reliability – related maintenance knowledge, such as Failure Modes, Effects, and Criticality Analysis (FMECA), or Fault Tree Analysis (FTA). To resolve the context of an event or interaction in a sociotechnical system, other types of context must be taken into account (Fig. 1). The detailed and domain specific modeling of the broad context categories may be further diversified depending on the exact nature of the application.

Context determination depends on linking data and knowledge and this is in line with the semantic web paradigm of linked data and product knowledge (Pistofidis, Emmanouilidis, Papadopoulos, & Botsaris, 2016). The real value of context understanding lies with the quality of data and knowledge upon which analysis, decisions, and actions are exercised. For example, (Fig. 2) raw data can be of little value if there is a lack of understanding about their provenance and underlying context. The figure illustrates that annotated data (information) can be more valuable if adequately analysed to obtain insights about the underlying generating processes, leading to knowledge-enhanced data, which are more likely to drive action recommendations. A sound understanding of the data context and overall situation awareness may produce additional insights and lead to more informed decisions or changes (e.g. supplier selection in Fig. 2). The enhanced value of product data across such a data value chain justifies the viewpoint that data is to be considered a value adding asset itself (Kubler et al., 2015). Context information management is a semantically scaled extension of information fusion for IoT (Snidaro, García, & Llinas, 2015), enabling links between human and non – human actors. Therefore, such links facilitate more effective information flows, better interfacing between actors, and allow for a more efficient integration of HIL in production and specifically PLM.

A Visual Analytics environment can be a significant enabler of such context-driven interaction. From the early years of exploratory data analysis (Tuckey, 1962) (Tuckey, 1977) all the way to current big data analytics (Idreos et al., 2015), the quality and value of data-driven decision making depends among other on data pre-processing initiated by human experts. Whereas computer data analysis in the past had a very limited set of options for

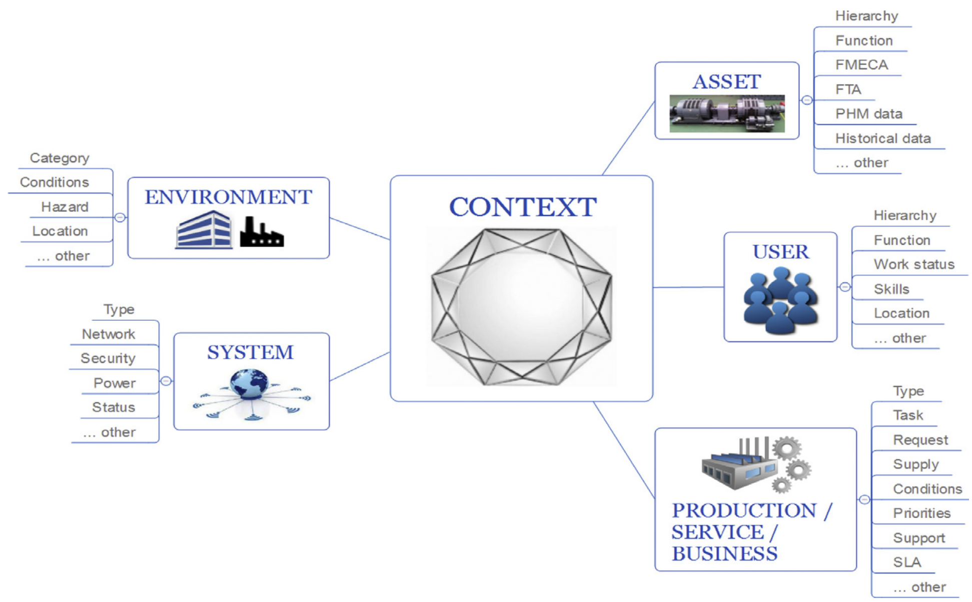

Fig. 1. Context categories in product and asset lifecycle management.

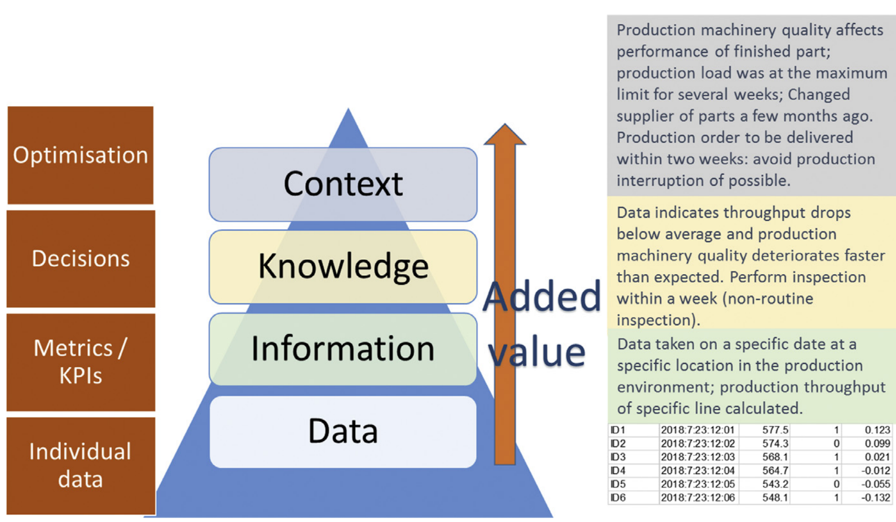

Fig. 2. Context and data value.

user interaction with data, current visual analytics greatly upgrade interaction capabilities, steering expert judgement through visually presented aspects of data characteristics (Thomas and Cook, 2005). Visually enhanced data presentation makes it easier for users to comprehend the data context compared to reviewing raw data (Endert et al., 2014). Situational awareness and context management also play a key role and can be combined with Visual Analytics, for example, aiming at combining automated processing through condition monitoring with human contributed observations, as a means for context-based information fusion for diagnostics (Emmanouilidis et al., 2016). Furthermore, situational awareness is invaluable in resolving context ambiguity in handling industrial alert and alarm management, which if unresolved can overwhelm human operators with unmanageable numbers of alerts (da Silva, Pereira, & Gotz, 2016).

The overview of related research highlights that there is much to be gained through the integration human and non-human actors in sociotechnical systems but for such integration to become more effective, further research is needed to align such actors not only by means of technology integration but also by appropriate designs for seamless data, information, knowledge, and decision flows. The next section introduces the concept of placing such designs within the viewpoint of cognitive capabilities in sociotechnical systems.

3. Cognitive capabilities in sociotechnical systems

Even with the integration of a range of Industry 4.0 technologies, production environments are still far below the level of intelligence normally associated with human actors. While non-human actors exhibit some level of intelligent function, the active presence of a human actor enables more powerful cognitive capabilities to be expressed in production activities. Such activities can be considered to draw parallels with the capabilities of cognitive architectures (Langley, Laird, & Rogers, 2009). While such capabilities have been studied regarding human cognitive abilities and artificial cognitive systems, the potential for joint sociotechnical systems, which both amplify human cognitive abilities and expand the capabilities of technical systems, have only recently gained attention (Arica et al., 2018). Activities which are closer to becoming industrial practice involve a subset of cognitive capabilities and are

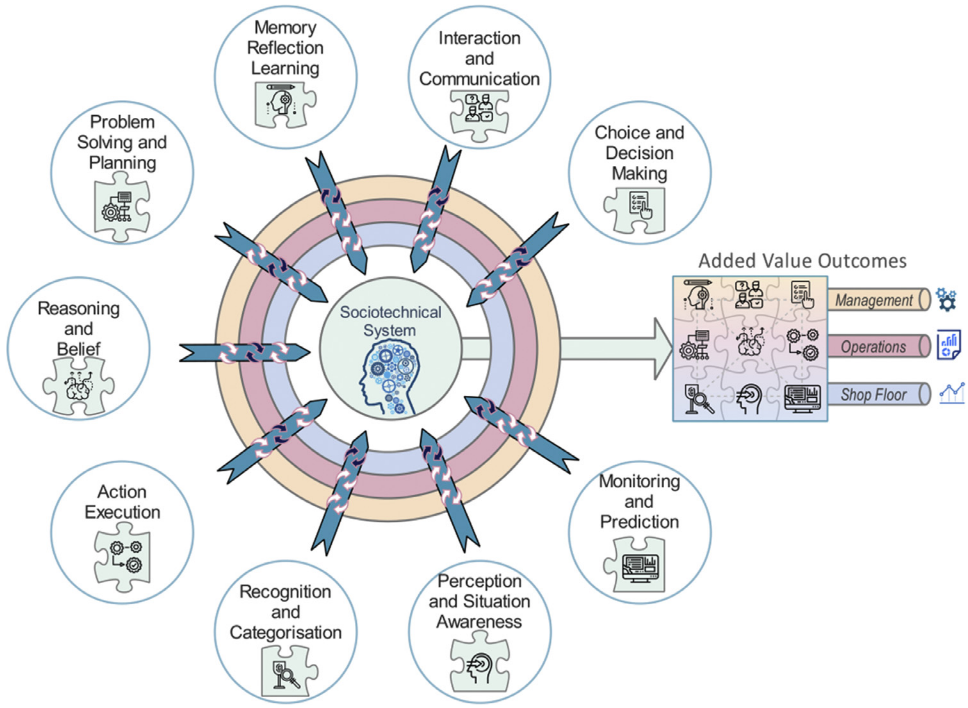

Fig. 3. HIL-CPS cognitive capabilities.

linked to sensing, state inference, action generation and execution in a form of HIL intelligence in CPS systems (Nunes et al., 2015).

This paper argues that the integration of such cognitive capabilities can also be exploited in industrial information processing cycles. Fig. 3 illustrates the concept of integrating cognitive abilities in sociotechnical systems to deliver added value outcomes for industrial environments.

While in the figure outcomes are highlighted for shop floor, operations, and management, HIL integration has broader potential in bringing the range of human cognitive capabilities in the whole production and product lifecycle management activities, joining the power of CPS entities with production, operational, and information technologies. For such a potential to be realised, a range of issues in integrating individual human capabilities within technical systems, still need to be resolved. For example, the lack of sufficiently flexible interaction interfaces, despite many advances at the perception and recognition level, is still a significant hurdle. Recent studies have shown the dominance of intuitive cognition over proper reasoning and decision making, a finding that calls for further research in human-automation interaction in industrial settings (Patterson, 2017). To this end, recent efforts focused on modelling human activities within CPS systems (Fantini, Pinzone, & Taisch, 2018), identifying specific challenges on: (a) understanding and controlling the interaction between workers and CPS entities; (b) how to capture the added value of human-contributed activity; and (c) how to take into account and match a specific situational context with skills and characteristics of workers. Context Information Management is a key concept aimed at situation awareness and is therefore appropriate for addressing the above highlighted challenges.

CPS as well as product and asset lifecycle management activities increasingly generate a very significant amount of data. While automated data analytics expectations are high, there are still many situations wherein placing the HIL of analytics is highly beneficial. Enhancing human cognitive capabilities in analysing data via relevant software tools is linked to the concept of "the human is the loop" (Endert et al., 2014), indicating the valuable role of human analysts in integrating their cognitive abilities when interacting with visual analytics environments. A conceptual view of HIL in the Visual Analytics process is introduced in Fig. 4.

In a typical visual analytics scenario, data are managed, aggregated, and retrieved through adequate data management business logic. Data visualisation options are offered through an analytics wizard, offering different data visualization and exploration options and retrieving the most relevant ones. The user can interact with the analytics and visualization tools to direct data processing, enrich data, process them to obtain prioritisation ranking, and in case of critical unexpected events being detected, to issue relevant alerts, or otherwise routinely present summarised information with relevant dashboards. This process augments part of the cognitive processing cycle presented earlier, including recognition and categorisation, interaction, perception and situation assessment, monitoring and prediction, as well as decision making.

Having introduced the key concepts of human cognitive capabilities in sociotechnical systems, highlighting their role in HIL-CPS and Human in the Visual Analytics loop, the next section presents the implementation of such concepts on selected application cases.

4. Human-in-the-loop in production environments

4.1. Introducing application cases for HIL in production environments

The role of HIL is highlighted in two application cases, both relevant to production environments. They employ condition monitoring and they involve some common design elements, but in a manner to serve different application needs. In that sense, they feature diverse context and implementation choices and they aim to enhance joint sociotechnical capabilities in distinct ways. The first case actively involves HIL for Linking Data and Knowledge Management and for enhancing machine learning capabilities in order to integrate operating stage technical systems (condition monitoring) with design stage knowledge (e.g. Failure Modes, Effects, and Criticality Analysis – FMECA). The second case engages HIL with visual analytics to communicate condition monitoring

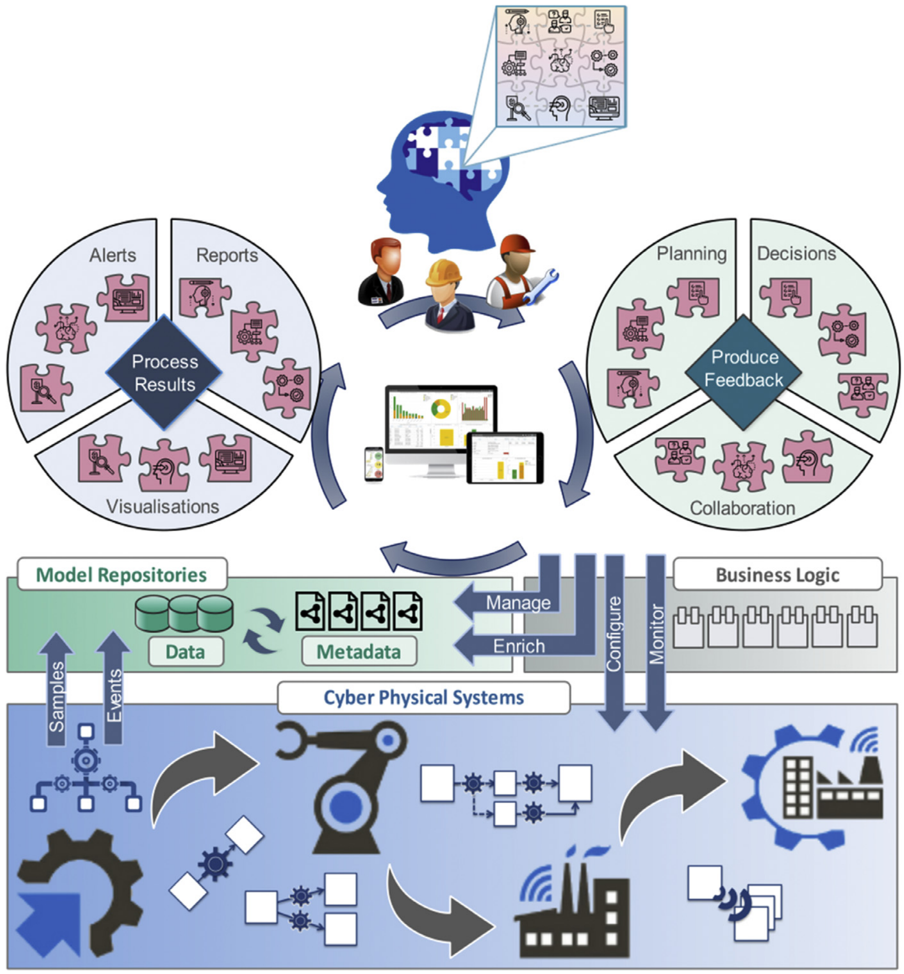

Fig. 4. Human in the visual analytics loop.

outcomes in visually relevant ways for augmenting human capabilities when performing analytics-driven decision tasks. The first case illustrates the amplification of machine learning and condition monitoring capabilities via HIL, while the second one is a case wherein human capabilities when interacting with technical systems are themselves enhanced. Even though the direction of the amplification between technical and human capabilities is different, the ultimate result in both examples is a joint sociotechnical system capabilities enhancement.

In both cases the technical system employs vibration condition monitoring to associate extracted vibration signal parameters to machine or component condition. The first case additionally employs machine learning to learn this association. The ability of machine learning to associate new patterns to conditions not yet covered by historical data or to bias learning on the basis of human expert knowledge is aided by human interaction. While blind data-driven machine learning might eventually learn such additional associations, this natural form of human intervention makes the learning process more focused. The extracted vibration features in both cases were as in Katsouros et al. (2015) and comprised the signal RMS, skewness, kurtosis, shape factor, crest factor, peak value and impulse factor. A stream of data acquired in this way is a timestamped sequence of vectors $x_i \in \mathbb{R}^7$, $x_i = [rms_i,\ sk_i, k_i,\ sf_i,\ cf_i,\ p_i,\ if_i]^T$ with $i = 1, \ldots, K$, wherein the vector parameters correspond to the extracted features from the measurement signal described above, and K denotes the number of samples in a data stream.

4.2. Linked data, knowledge, and machine learning

The data and knowledge processing chain in maintenance practice seeks to upgrade the added value of collected data to drive more effective evidence driven decision making. Event detection, diagnostics, prognostics, and decision support for maintenance have received much attention, but one of the areas where further research is needed relates to the way such solutions can enable and indeed benefit from HIL interaction. In this example, HIL concepts are employed to fuse automated data processing with human contributed knowledge in maintenance decision support (Pistofidis et al., 2016). Evidence driven decision making in this way is not solely data-driven or based on prerecorded knowledge but is enriched with mechanisms that bring together human and non-human actors in a way that links data, knowledge and machine learning (Fig. 5).

The application environment is that of an industrial production facility for lifts and an actual in-service operating installation. This is a representative case of a hydraulic lift that has a very wide base of residential and office installations. Preventive maintenance is performed on a monthly basis to detect failure events and drive recommended maintenance actions. Vibration sensors are positioned in two areas of interest in a lift installation, namely, at the bearing of the lift's drive motor and at the rollers of the lift's cabin. The experiments focus on analysing the signals obtained from the cabin rolling wheels to infer asset condition so as to drive maintenance actions recommendations. While automated approaches

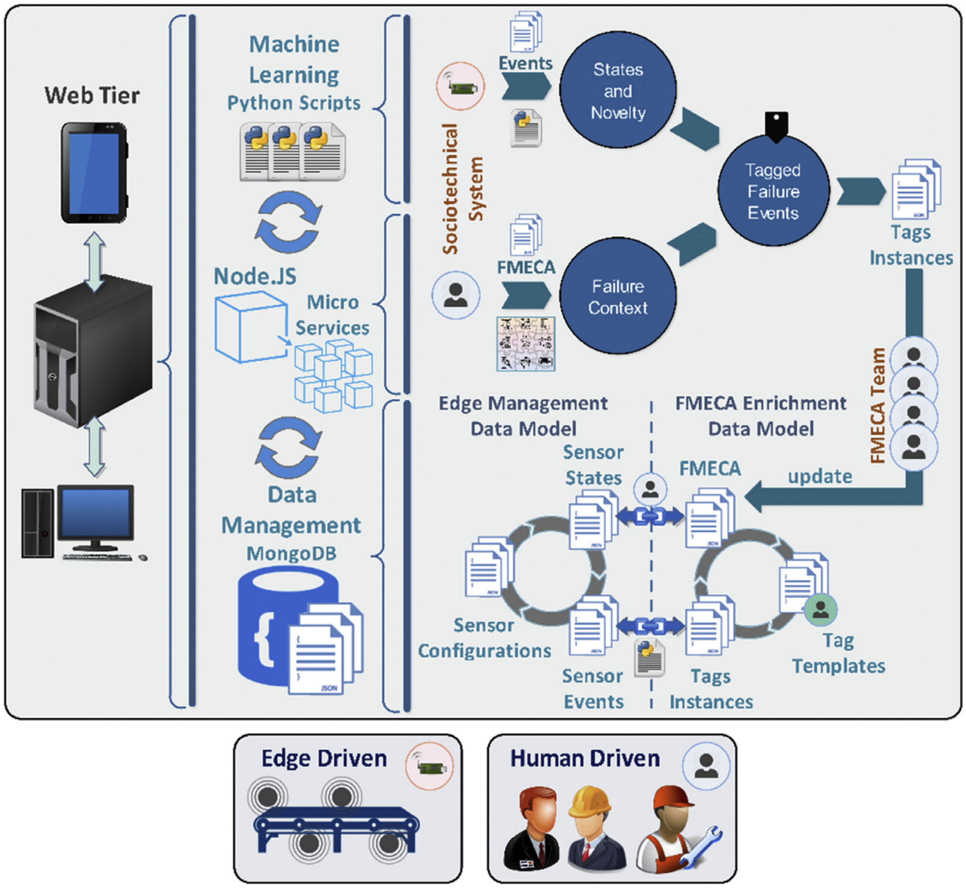

Fig. 5. Architecture for Linked data, knowledge, and machine learning.

may succeed where data are sufficiently representative of the underlying knowledge, the proposed approach mitigates risks when this is not the case. The approach is demonstrated through an e-maintenance platform, comprising components for machine learning, embedded event detection and diagnostics, as well as linked maintenance knowledge management and action recommendation (Pistofidis et al., 2012).

The objective of this case study is to show how integrating automated processing and human-contributed observations can be fused with established reliability maintenance constructs, such as FMECA, to expand the capabilities of either human actors or technical monitoring systems operating in isolation via a joint sociotechnical solution (Fig. 6). Furthermore, it extends machine learning capabilities by allowing HIL to contribute to the learning process, enabling it to be driven not only by data but also by integrating human expert knowledge.

The design approach to HIL in this case is two-fold. First it aims at facilitating maintenance knowledge management, including sharing, enrichment, validation and extension of knowledge. Second, it aims at a monitoring abstraction which can be employed with machine learning and HIL to deliver customisable condition monitoring implementations able to learn not only from data but also from HIL involvement. This is presented in more detail in the next sections.

4.2.1. HIL for knowledge enrichment in maintenance within production

A key addressed challenge is to design a metadata management system (MMS) that binds the semantics of sensor events with reference diagnostics for failure modes. The linking process includes human experts in the loop, utilising and progressively enriching the backbone of an FMECA study with input from both the technical system as well as maintenance personnel. Originating from the smart sensors that populate the edge IoT tier of the architecture (Fig. 7), sensor events represent an asynchronous flow of information that signals events of interest, which may correspond to the identification of known and novel states.

These events are handled by services that examine whether failure events are currently linked with the triggered states. Appropriate input is then produced through the use of semantic annotations that characterise failure events as either ‘confirmed’ or ‘False Alarms’. Human experts are alerted to intervene when novel events or events linked to failure modes are detected. Both processes, ie automated and human-triggered data tagging, produce semantic data annotations, ie metadata. Overall, the produced timelines of semantic annotations is collaboratively driven jointly by smart sensors and human experts (Fig. 5), but the latter occur at a different time scale compared to automated alerts. Risk quantification for monitored machinery and infrastructure is pursued through Risk Priority Number (RPN) evaluation, as in a typical FMECA study (RPN = Severity X Occurrence X Detection) but the estimation is now influenced by the joint action of the technical and human actors of a sociotechnical system.

The methodology encourages human involvement to provide direct feedback and event validation. This is achieved by simplifying human feedback, empowering staff to participate in knowledge flows with minimal, intuitive and natural interfaces. Such a process is more likely to secure staff participation, offering a more familiar input pattern that has been for years now driving the analytics of enterprise social networks (tags, ‘likes’, and short messages). It is a pattern that seeks to aggregate a large volume of concise inputs into a knowledge building process that invests in collecting

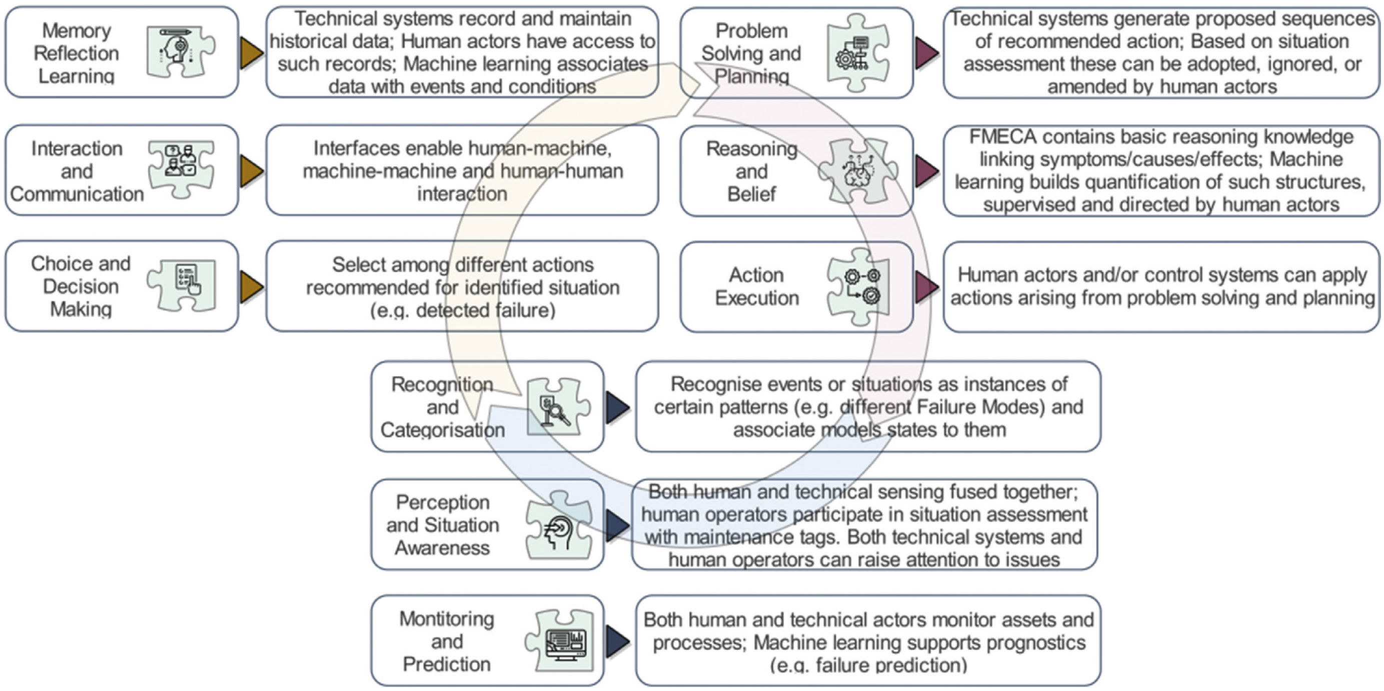

Fig. 6. HIL and cognitive capabilities in machine learning for decision making.

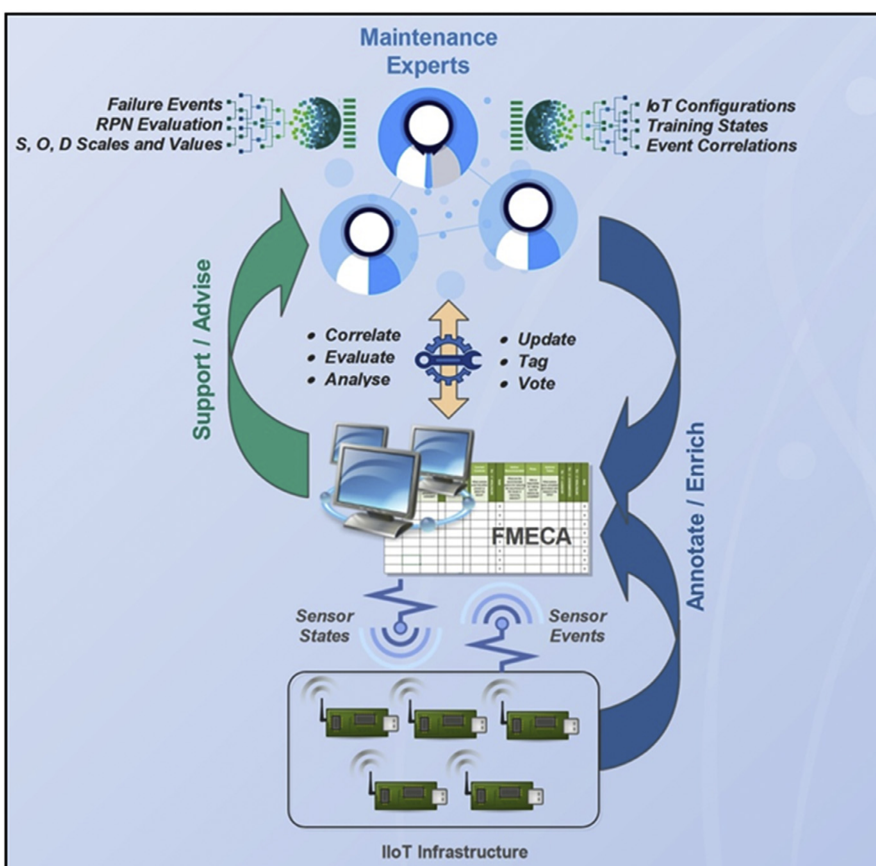

Fig. 7. Functional view of the design for human in the loop knowledge enrichment.

swift evaluations over the validity and quality of given information. In this approach, it is less important to provide a higher quality starting knowledge (initial version of an FMECA study), and more important to provide the context, the semantics, and the tools to create a virtual collaboration infrastructure to assess its validity, acceptance and connection to practice. It is exactly this collaboration, driven by short swift evaluations (maintenance tags), that allows the human to actively contribute in a knowledge validation loop. The FMECA study is the reference knowledge baseline. But it is not static and its enrichment, evolution, and validation can be triggered as a product of consensus and fusion from a larger group of technical staff and experts. To secure the quality of this input and furthermore manage the complexity of multiple similar contributions, user roles and votes are employed. Only specific roles can facilitate tags and votes, including experts that participate in the FMECA review team and maintenance engineers with strictly defined function domain. Every contribution is logged and when reviewed (FEMCA review) should be backed with evidence (condition monitoring history). This FMECA update comprises corrections, proposed extensions, and builds the enriched evidence to back them up. It is not automatic but is triggered by the joint action of the technical and human actors.

The design approach in this application case delivers a sociotechnical tool that provides access to services that offer: (i) a shop-floor contextualised FMECA study; (ii) a streamlined review collection process; (iii) and a filtered view of human and non-human actors' annotations. Current IoT and cloud – oriented patterns and technologies (micro-services, Node.JS, MongoDB) were employed in the tool design and implementation, but two specific design aspects are highlighted next, as they define the two key perspectives of how HIL functions in this application case are: (a) supervising and tracking the enrichment and versioning of knowledge and (b) encouraging personnel to collaborate and share experience and critical thinking over working practice experience and established knowledge.

4.2.1.1. Data provenance. A timeline of events can reveal patterns that impact on maintenance diagnostics and risk assessment. Each entry in a maintenance system is time-stamped and all actions are logged. Data Provenance refers to the ability to trace and verify data creation, and, in our case, failure evaluations and important sensor events, in the form of maintenance metadata. Provenance of such assessments can identify patterns that depict asset reliability and may offer hints for risk analysis. Therefore, adding a metadata layer on top of sensor and reliability data further enhances provenance by collecting the evidence of a validation loop. Populating this layer and driving this loop, human observations and machine generated events, produce metadata, adding background knowledge and evidence to support an FMECA review.

4.2.1.2. Context sharing. Social networks or social-network inspired features are increasingly included in enterprise communities and collaborative environments. Professionals are becoming familiar with such features in standard social networking context and can benefit from their inclusion in enterprise environments, allowing them to offer their input or annotate the input of others, at a real-time manner and with many sharing options. More specifically, information and features such as 'voters' and 'votes' stimulate virtual interaction and conduct a social contextualisation of shared content. Instead of long forms for reporting, the proposed methodology employs minimal input via customisable maintenance tags with voting options. This is no substitute for FMECA revision decisions, but allows the accumulation of evidence and capturing of observations and knowledge from personnel. If not enough metadata are clustered around a specific failure event, a single observation or sensor event is less likely to trigger the appropriate

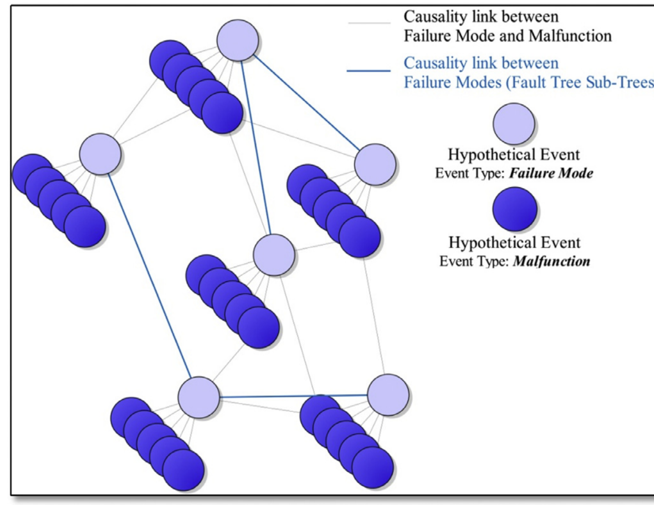

Fig. 8. A causal semantic graph of events.

re-evaluation during the FMECA review process. Votes are an extra context sharing feature that enables personnel to actively contribute to a crowdsourcing and sharing of observations, evaluations, and sensor events.

Comprehending the role of metadata and annotations as units of risk-oriented maintenance knowledge can be more effective with a tool that manages them on top of a widely accepted knowledge backbone, namely the FMECA study. The swift capturing and sharing of such maintenance knowledge by personnel, together with data and events produced by the technical system, leads to an incrementally enriched version of FMECA. This is a principle similar to crowdsourcing intelligence in recommender systems, whereby the users' collaborative contribution is exploited. Leveraging upon one of the most significant enterprise assets, namely the human factor, is the key to facilitating more effective knowledge flows within the enterprise.

With increasing adoption, Linked Data have introduced formalisation frameworks and technologies that can efficiently instantiate knowledge representations. Such frameworks can employ annotation for tagging important content. Supporting technologies provide metadata contextualisation using widely established data constructs. The amalgamation of relevant knowledge, data, and annotations define the Failure Context as the confluence of factors contributing to the occurrence of a failure. In other words, the Failure Context holds the combined knowledge relevant to the occurrence of a failure mode and the assets-specific time-relevant feedback of maintenance practice. Building upon the established semantics of an FMECA study, this design provides the means to formalising and instantiating this context. Adopting structures and components from the established MIMOSA schema,[1] the FMECA model is customised to empower the creation of an event map (Fig. 8).

This map is a semantic graph where Failure Events act as the core nodes. Implemented as a distinct set of semantics, Maintenance Tags are used to annotate core or supporting nodes and thus report why, how and when a failure event occurs. The initial version of this graph is created as the product of the very first FMECA study completed by the appropriate team of experts. As dictated by common practice in reliability engineering and risk analysis, an FMECA study is followed by scheduled reviews and evaluations. The sociotechnical system approach introduced in this paper enables enhanced FMECA review and evolution, through feeding into

[1] www.mimosa.org.

it collective evidence from the technical and human system actors. Information and knowledge about failure modes become increasingly important, when human input (from maintenance practice) and technical triggers (from the machine learning – enabled condition monitoring system) repeatedly validate known and reveal new connections (causality links) between them and other events and failure modes. The collected input is a distinct second layer of knowledge, above and directly linked to FMECA (metadata). FMECA reviews continue to be the milestones in the enrichment process, but they can now benefit or even be triggered through this second layer, and drive insights for the corrections and extension of the next FMECA version, which remains the responsibility of the FMECA team.

The introduced model utilises MIMOSA as a starting point for drawing a subset of its core semantics. Specifically, in MIMOSA, failure events are denoted with the entity *Hypothetical Event*. In the new model, the *Hypothetical Event* profile has been extended with attributes that record semantics of occurrence and detectability. These hold a ranked evaluation for the events' frequency and detection probability. Along with the inherited property for event's severity, these scaled attributes can drive a RPN-based evaluation. Furthermore, acting as the building block of a growing event map, the introduced event entity capitalizes on distinct and well-defined recursive attributes that link events with cause and effect associations. A "malfunction" and a "failure mode" are both types of the event entity. Failure modes have populated causes, effects and solutions that are associated with them. A malfunction is primarily a simple event linked to a failure mode. This is a process that gradually builds a knowledge infrastructure for an Asset Fault Tree. Traversing such event links may facilitate a root cause analysis and provide insights for risk assessment.

Whenever a user is prompt for an assessment, it is practical to provide a starting reference point, FMECA knowledge in this case. In the present design the referencing dynamics of '*maintenance tags*' are employed to create metadata that bind users' feedback with FMECA knowledge. Instead of the simple string tags, commonly used in the context of semantic web, the present design employed class-types of review assessments, which acquire added value when coupled with FMECA content. Each tag has a straightforward use and annotation purpose that is defined by its tag template. The default set of tag templates is configurable and extendable. Only specific roles (e.g. roles with FMECA review authorization/'*facilitator*') can create, modify and adjust the type and purpose of tag templates. The ability to extend and map the enrichment process often resides in the skillset of experts that clearly understand the scope, depth and purpose of the annotated knowledge. An experienced engineer can be trained to translate new maintenance goals and policies into meaningful tags that create new actionable semantic links.

A tag instance is the modeling entity for maintenance metadata. Every annotation action creates a tag instance. Each instance constitutes a timestamped unit of maintenance knowledge. Tag instances are shared and can be searched or filtered by users. Their knowledge can be further enriched with tag votes (declaring agreement) and tag mini-forms (additional feedback). These are termed as Maintenance *micro-Knowledge* (Fig. 9).

The enriched version of FMECA is essentially the content that is dynamically tagged and identified as contextually relevant to how real events manifested and occurred, and why specific maintenance solutions or diagnostic interpretations were reached. In this sociotechnical system design, maintenance tags can be produced by both human and non-human actors. Incorporating these two information flows with an FMECA study by translating them into brief and accurate review annotations constitutes a simple enrichment process that formulates a growing pool of maintenance metadata that is natively organised and collaboratively evaluated.

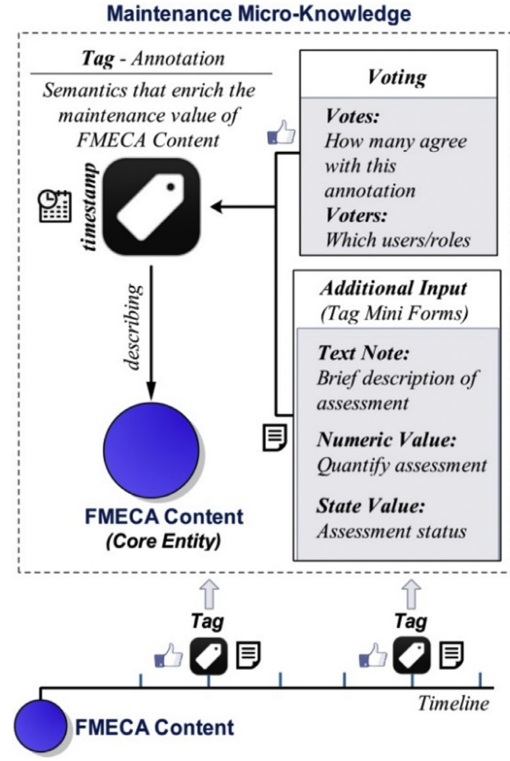

Fig. 9. Maintenance micro - knowledge.

This fused information pool containing enriched failure mode profiles and timelines of maintenance tags, can be consumed and analysed for mismatches, corrections and additions to be inserted in the next version of FMECA, or help document and support critical risk assessments and maintenance plans.

The metadata are instantiated and stored in a document-oriented database (Fig. 5), allowing the creation of structured, semi-structured, unstructured and polymorphic data. Its ability to handle and query massive volumes of new and rapidly changing data types meets the design decision to allow the creation of custom tags and encourage the collection of more and better organized human input. Furthermore, the implementation of the back-end logic is consistent with micro-services patterns, breaking the application logic into smaller modules, enabling better flexibility and laying the groundwork for cloud compatibility. This application case implements the concept of capturing the tacit knowledge relevant to maintenance practice and risk assessment, while also acting and planning upon maintenance events that can support maintenance intelligence. It captures and transforms maintenance expertise into knowledge fragments that instantly link background reliability knowledge to everyday practice, and crowdsources data, information and knowledge from human and technical actors.

4.2.2. HIL in machine learning loop for asset self-awareness

Machine learning is typically considered as an automated process driving decision making. However, data-driven learning is often inefficient in complex tasks with poorly representative data. Incorporating HIL in machine learning is rarely considered although it can make a real impact on real world applications, such as in production environments. The design of the HIL solution in the present application case considers a machine learning infrastructure distributed among the edge and the web service level (Fig. 5). The web service level involves operations needed for learning to model the associations between signal features and asset conditions and the management of all relevant operations to manage

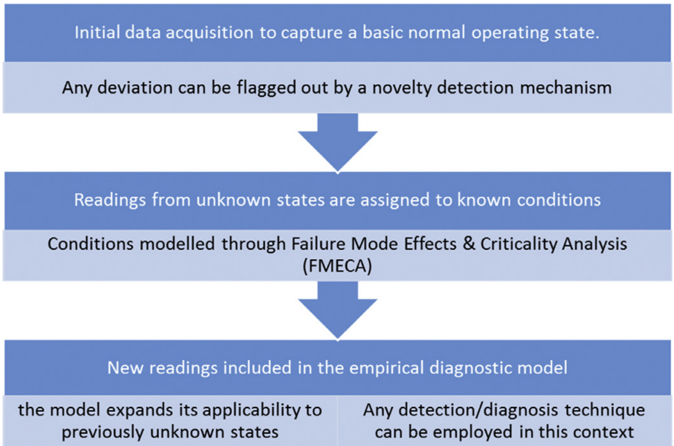

Fig. 10. Incremental learning of process states.

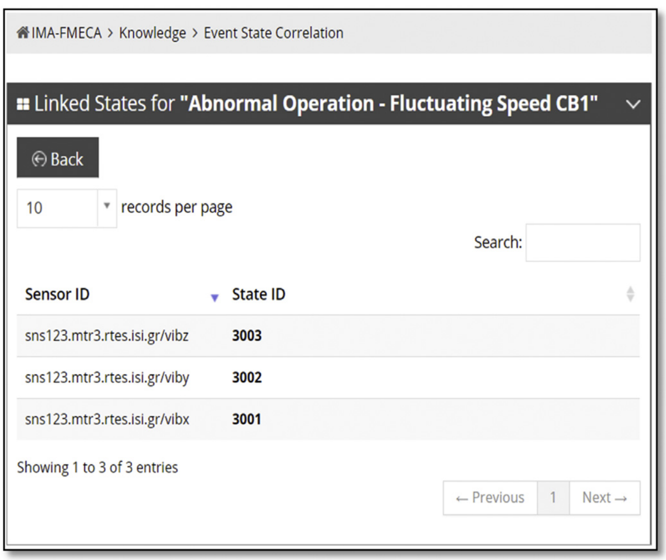

Fig. 11. Association of failure modes with Gaussian states.

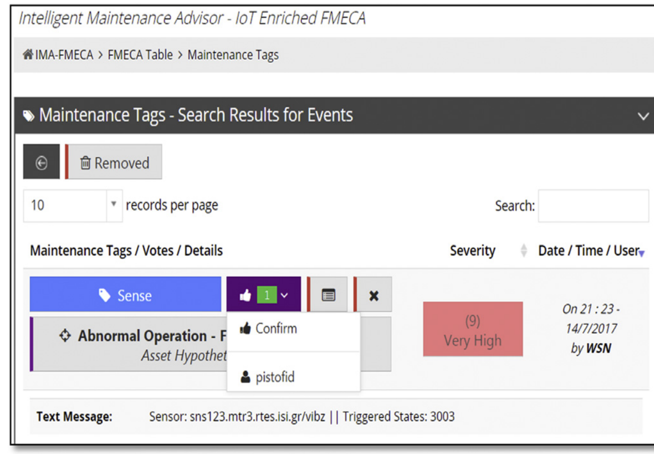

Fig. 12. Expert confirmation of sensor driven annotation.

such processes and communicate results. These operations employ machine learning, inclusive off a human feedback loop for states detected as novel or for validation (Fig. 10). The edge implementation includes sensor data acquisition, feature extraction, as well as novelty detection and classification of sensed data to asset conditions and is described in more detail in the next section.

Numerous approaches have been developed to model changes in system states, identify current states, and predict future ones for different prognostics purposes (Lee at al., 2014). The implementation in the present research adopted the parametric family of Gaussian Mixture Models (GMM) for modelling the failure modes. GMMs belong to the parametric family of multivariate Gaussian density function, which can be combined with Bayesian statistics to make statistical inferences regarding the failure modes.

GMMs can be utilised to approximate an arbitrary distribution within an arbitrary accuracy. For the Gaussian density function, it suffices to calculate from the feature data points the mean vector and the covariance matrix. In order to keep the implementation simple, the choice made is to relate each failure mode to a number of Gaussian components/states that result from the training data collected during the operation of the system. Fig. 11 shows a user interface from the implementation example where a user can effect the above association. This view displays the association of a failure mode with a set of model states from specific sensors. This association is performed off line and failure events can be profiled by experts and linked with states. Sensor events from the edge node, can be processed and failure events linked to the triggered states are tagged and confirmed. Such sensor events and human input produce metadata and can serve as example patterns for machine learning. Rather than relying on blind data-driven only learning, this effecticlety constitutes an incremental, HIL-enabled learning process.

For the calculation of the covariance matrix the Minimum Covariance Determinant (MCD) estimator is employed, which is among the robust estimators of a data set's covariance (Rousseeuw & Leroy, 1987). In the experimental setting the feature vectors are sets of statistical time series parameters calculated on shifted windows over the sensorial data stream (Katsouros et al., 2015). The bootstrap of the system is based on a model of the normal operation mode of each asset, which is trained from feature sequences that have been collected from the sensor level. If there is prior knowledge of sensorial features' association with failure modes then such associations can also be included in the initialisation of the system. The parameters of the Gaussian models are communicated down to the sensor level. The embedded algorithm at the edge node calculates the feature sequences and their degree of classification to each of the known states/modes. Features that cannot be classified to any of the known states are marked as novel.

For states which are defined to be determined at the sensor node, the embedded classifier assigns readings to states for non-novel data. This is done by calculating an overall degree of classification, applying a Bayesian approach using the independence assumption for the feature data points and weighting the product of the probabilities with the a priori probability of each state. For more complex states that require data from multiple sensor nodes, the classification of the feature sequences is communicated to the relevant web service, which assigns readings to classes in a similar manner, but for multiple feature sets from different nodes. The feature sequences that belong to states of failure modes will trigger alerts, drawing attention for human intervention. In the case where the event is related to an existing failure mode, the web services use maintenance tags to report it and the human expert may verify the event or raise doubts about it, potentially as a false alarm not related to the failure mode (Fig. 12).

Maintenance events or alerts can be issued by technical (non-human) or human actors (Fig. 5). Alerts related to detected novel events correspond to events that are not classified in any of the known states, ie normal conditions or failure modes for which representative data are available. Such novel data are stored in order to be examined at a later stage by a human expert. These may represent a new state that is extending one of the known conditions. Alternatively, they can be related to a new failure mode. In the latter case the human expert has to update the knowledge base with the new failure mode and relate this mode to the relevant sample data, so as Gaussian models can be built on the basis of the related feature sequences. This is an example of machine learning wherein human input is a significant part of the learning process, playing a role that is very hard to be replicated by automated machine learning. Such HIL in the machine learning loop can therefore have a positive impact on enhancing data value chains from the edge to the cloud. The next section presents an edge implementation that benefits from and makes use of HIL in machine learning.

4.2.3. HIL for IoT-driven event detection

The edge implementation for IoT-driven event detection is directly benefitting from the HIL of the machine learning process described in the previous section, as the human interaction enables the association of sensed data and features to states. State models are then downloaded to the embedded device to drive edge analytics for event detection and diagnostics. The parts of the machine learning for novelty detection and diagnostics that have been assigned to execute at the edge of the system, close to the monitored resources, were implemented using a range of embedded devices of different capabilities, interconnected under a wireless sensor network of heterogeneous IoT technologies. HIL contributed knowledge, such as the association of failure modes and machine states is effectively driving the embedded node operation. The overall implementation was integrated within an e-maintenance platform (Pistofidis, Koulamas, Karampatzakis, Papathanassiou, & Centre, 2012), while the algorithmic implementation described in the previous section (Katsouros et al., 2015) can trigger production process adaptation driven by IoT-captured failure events (Alexakos, Anagnostopoulos, Fournaris, Kalogeras, & Koulamas, 2017).

The implementation over diverse hardware platforms and technologies is indicative of the wide range abstraction possibilities, making the approach appropriate for serving very diverse requirements. The versatility of the approach is also evident via the application layer connectivity, which is offered through RESTful Web Services,[2] implemented using JSON[3] over HTTP[4] or CBOR[5] over CoAP.[6] These services allow accessing the sensor nodes directly or through the appropriate gateways, depending on the sensor node capabilities. All sensor nodes execute sampling, storage, feature extraction, classification and novelty detection services, with their operation externally controlled through a well-structured set of resource URIs, accessed through HTTP or CoAP GET/POST/PUT REST operations. Key resources include:

- /info: This is a set of implementation dependent non configurable parameters (basic buffer size, maximum number of buffers, the maximum number of sliding windows for the acquired signal processing, supported sampling rates and maximum number of supported state descriptions).
- /config: The novelty detection engine configurable parameters (sampling rate, sliding windows size and step, sample scaling factor, monitoring period).
- /stateset: A set of configurable parameters, ie state descriptions and thresholds over which the distances of the calculated feature matrices are compared in order to generate a novelty event; these can be ‘learned’ following the previous section’ process and downloaded to edge device.
- /event: This is the observable resource of the last event triggered by the node, modelled through a structure that is sent to the upper software layers, periodically, on-demand or when an alert is triggered by the novelty detection mechanism. It encapsulates the whole feature-extraction, novelty detection, and classification process chain. The novelty detection implementation can instantiate and selectively return either only a Boolean result of a novel state, that is a calculation window with features distance higher than the configured threshold from a known state; or, additionally, all feature values for the window and the distances from all known states, quantifying dissimilarity.

According to the application requirements, there are different measurement classes in terms of sensing elements and relevant qualities. Low-end specification correspond to simple scalar values from temperature sensors with limited requirements for sampling bandwidth and processing. These can be sufficiently supported by resource constraint embedded devices and low bandwidth protocols for IoT networking. At the high-end there are high quality, multiaxial vibration monitoring requirements that may pose higher power and processing resources requirements, so as to manage more complex data streams and interfacing to industrial grade IEPE sensors. In between, there can be mid-range nodes that can still handle series and vector measurements but with lower sampling rates and accuracy needs. These nodes may be supported by mid-range hardware regarding analogue and digital data acquisition, processing power, flash and RAM space for data acquisition, processing, storage and radio communication transmission. The edge design abstracts and supports all the above requirements specifications and a broad range of heterogeneous sensor nodes was implemented, covering the whole capabilities spectrum.

Specifically, at the lowest end, the embedded detection engine was realized with off-the-self hardware components and widely used IoT operating systems such as the TelosB/TinyOS platform, the NXP Jennic platform over its own API and over the Contiki OS, as well as the PrismaSense development kit platform and API, using ZigBee, 6LoWPAN and raw IEEE802.15.4 protocol stacks. In these cases, the REST resources were accessed through custom gateways translating JSON/HTTP requests into binary commands transferred over the aforementioned wireless protocols, as device capabilities could not support complex application layer protocols. At the mid-range level, a special resource constraint embedded node was created based on a two processor board system structure, separating the application from the communication processor, coupled with a number of exchangeable sensor interface boards.

The sampling, signal processing and detection components execute on a Freescale FRDMK64F embedded board, while a full IoT networking stack, based on an IEEE-802.15.4 wireless interface and the IPv6, RPL[7] CoAP and CBOR components, is implemented on the CC2538 based Openmote board (Fig. 13). These nodes can be accessed either directly through CoAP or indirectly though the HTTP/CoAP gateway, according to the resource URL structure in the configuration database, separating the access (“http://” or “coap://”), node FQDN[8] and resource path. Finally, the high end vector sampling requirements and the support of industrial grade IEPE vibration sensors have been covered by a sensor node based on the DT9837B USB acquisition system from Data Translation, controlled by an embedded PC or tablet device which provides

[2] https://en.wikipedia.org/wiki/Representational_state_transfer.
[3] https://en.wikipedia.org/wiki/JSON.
[4] https://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol.
[5] https://en.wikipedia.org/wiki/CBOR.
[6] https://en.wikipedia.org/wiki/Constrained_Application_Protocol.
[7] https://tools.ietf.org/html/rfc6550.
[8] https://en.wikipedia.org/wiki/Fully_qualified_domain_name.

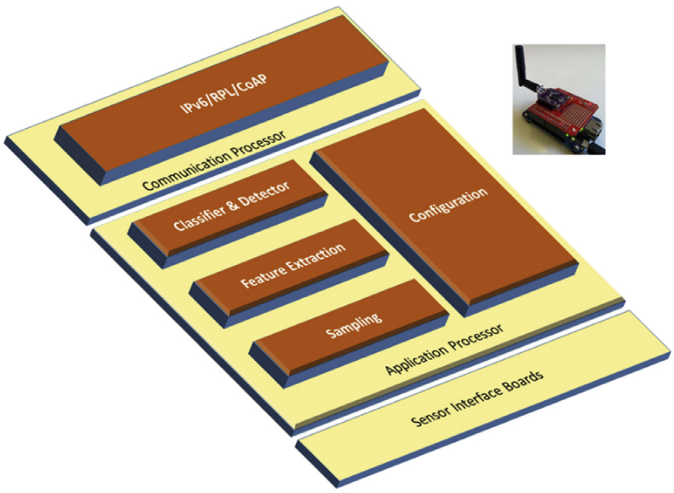

Fig. 13. Mid-range Embedded vibration monitoring with event detection capabilities.

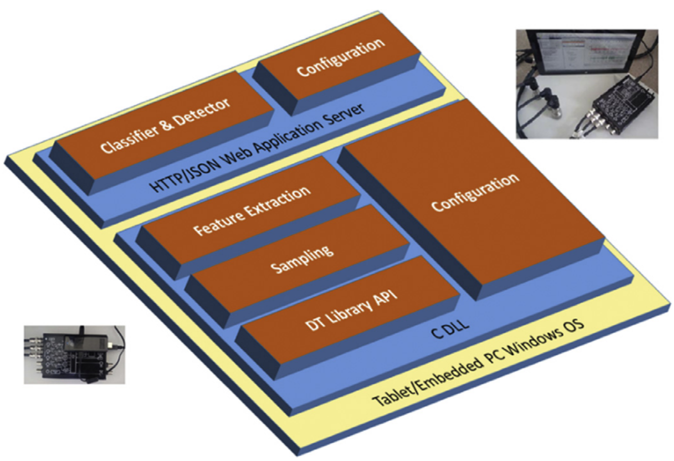

Fig. 14. High-end embedded vibration monitoring node with user interface.

the developed interfaces to the IIoT infrastructure, directly using HTTP/JSON (Fig. 14).

This IoT edge node implementation enables a monitoring asset to exhibit self-awareness features and issue event alerts when the processed sensor readings are identified either as novel or as belonging to failure mode states. It is part of a monitoring infrastructure that includes HIL in the definition or validation of the association of reference readings with asset conditions. This HIL interaction leads to an incremental machine learning approach, gradually building a more complete learned model over time. While part of this application example focused on HIL in machine learning for asset self-awareness and monitoring, the next application case utilises monitoring in a different context, linking maintenance processes in a PLM context with visual analytics.

4.3. IoT enabled visual analytics for linked maintenance and PLM

PLM tools already offer context-adapted product views enabling a user to view product representations, data, and information relevant to different lifecycle phase activities. However, user interaction in the data analysis loop could further benefit from due consideration of human cognitive capabilities to ensure a close integration of HIL in PLM activities. Visual analytics embed visual semantics in data representations, thereby employing simple but cognitively powerful means to better engage the human cognitive capabilities. In PLM activities, a user does not simply need to share product lifecycle data, but would benefit from doing so via visually enriched product views, which is the focus of the application case presented in this section.

4.3.1. HIL for IoT driven visual analytics

Considering that the most user-friendly product representation is a 3D product model, the key idea is to employ such a design – stage product representation together with MoL product information, related to product condition monitoring. By superimposing MoL relevant product information to BoL product views, such as a 3D product CAD model, linked maintenance data and knowledge (Pistofidis et al., 2016) become visual features of a product design representation, facilitating a user's understanding of MoL concepts, such as the occurrence of failure modes, within a design viewpoint. Therefore, this 3D visualization becomes a natural extension of standard analytics for monitoring data, including graphs of sensor readings and signal features, such as time domain and spectral features. This concept of blended digital product visual analytics was applied to design a laboratory based demonstrator for IoT-driven visual analytics. The demonstrator was developed on a mechanical transmission rig, comprising a lower shaft with four 42-tooth gears, driven by a motor, and an upper shaft with one larger 62-tooth gear, which is driven by the first shaft through meshing the upper shaft gear with any of the lower shaft gears (Fig. 15). Loading conditions can be adjusted with a brake, attached to the upper shaft, while the rotational speed is controlled by adjusting the motor speed. The lower shaft gears are initially identical but defects are introduced to gears 1–3, while keeping one gear in normal condition for reference. The defects are intended to emulate pitting, growing from smaller scale on gear 1 to a level consistent with extensive spalling, causing tooth pieces to fall apart (Fig. 16).

The aim was to produce an instantiation of the concept of linked knowledge in maintenance and PLM, with knowledge superposition to product views. The rig was retrofitted with a simple and inexpensive IoT monitoring arrangement and a software demonstrator was developed to offer visual analytics features. The aim is to highlight some of the possibilities for the amplification of cognitive abilities, which can be pursued by integrating this type of visual analytics with more conventional condition monitoring and PLM activities, as summarised in Fig. 17. For example, interaction and communication capabilities are offered through visual interfaces. Monitoring outcomes are communicated both via a 3D asset representation as well as standard signal graphs. The

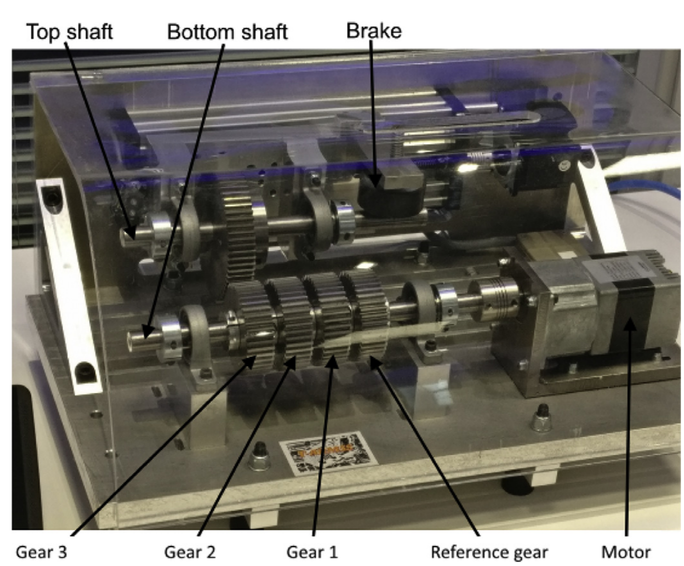

Fig. 15. Gearbox test rig.

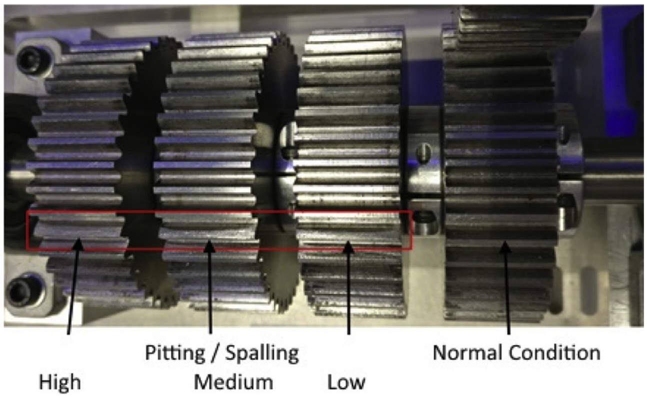

Fig. 16. Defect introduction.

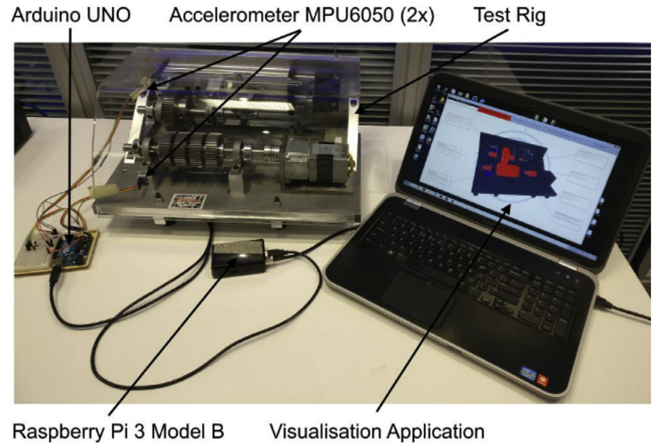

Fig. 18. Experimental setup arrangement on the gearbox test rig.

understanding of the situation by the user is aided not only by the conventional visual representations of monitored signals but also by superimposing the outcome of a detection mechanism to a 3D product representation and by highlighting in different colour parts that are considered to be developing a fault condition and therefore should be subject to further attention. This makes interaction and focus of attention more intuitive for the user.

4.3.2. Demonstrator of IoT-driven visual analytics

A low cost IoT – enabled monitoring solution, implementing data acquisition and basic diagnostics, was introduced for this application case. Rather than developing a thorough engineered solution, the demonstration objective focused on instantiating the basic data process chain for the blended digital product visual analytics concept. This process chain comprises

- data generation process, via a prototype data acquisition.
- a data processing stage, wherein acquired data are converted to monitoring parameters.
- a basic diagnostic stage, wherein acquired parameters are translated into asset conditions.
- blended visual analytics, jointly handling MoL data (e.g. diagnostics) with BoL (3D product model) product views.

The data generation process was implemented through an Arduino UNO board and two MPU 6050 accelerometers to capture gearbox vibration (Fig. 18). While this is not a sufficient set up for an industrially relevant solution, it is adequate for demonstrating the proposed concept and was selected for this purpose. The data processing stage was implemented on a Raspberry Pi 3 Model B board on Python, employing the SciPy library. This included signal averaging and extraction of standard statistical parameters from the acceleration signal as mentioned in Section 4.1, forming a sequence of measurement vectors. A Fast Fourier Transform (FFT) representation of the vibration signal is also calculated on board, after adequate filtering and windowing. The focus in this example is not specifically on the condition monitoring functionality but on incorporating the diagnostic outputs in an environment offering a visual analytics view of the product. Any other monitoring setup can be incorporated instead.

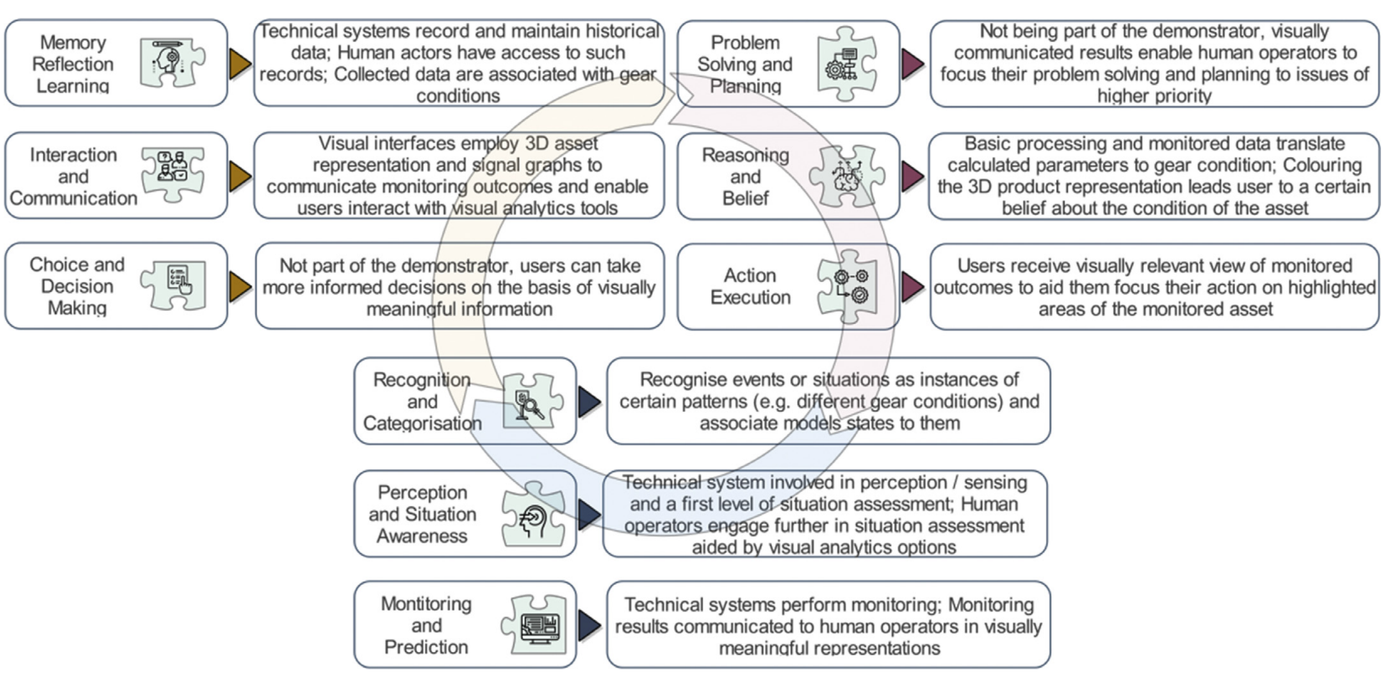

Fig. 17. HIL in visual analytics for condition monitoring linked with PLM.

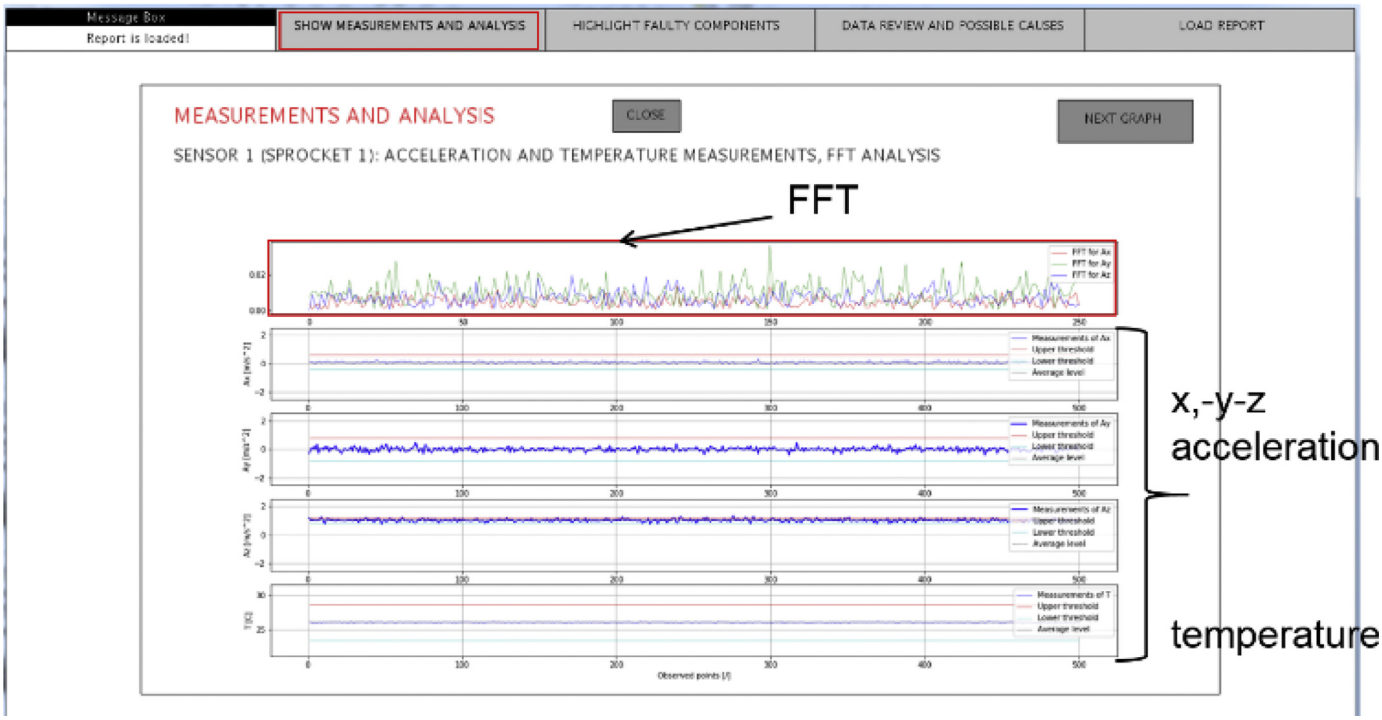

Fig. 19. Visual analytics example from the demo application.

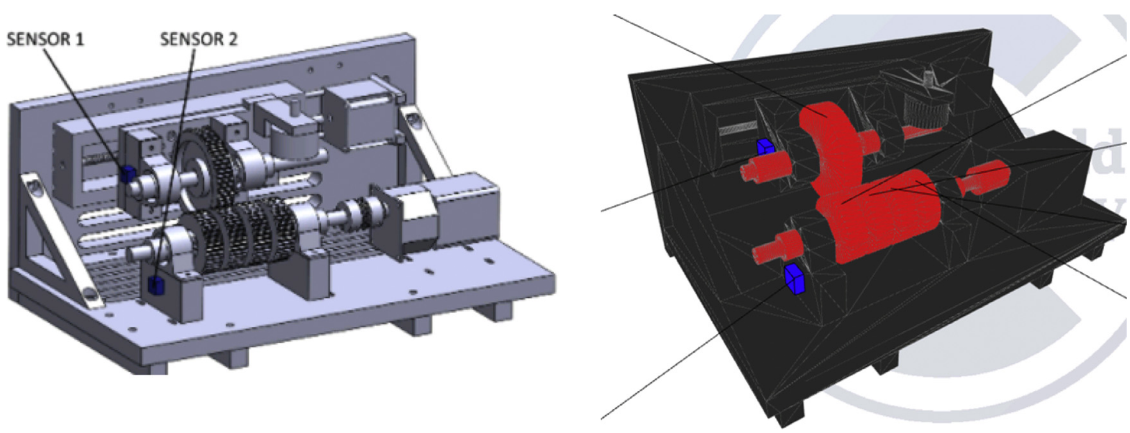

Fig. 20. Visual communication of measurement locations and diagnosed failure modes.

Reference data acquisition experiments were performed with different gear coupling setups, starting with gear without defects to obtain reference data from normal operating condition. Data were acquired with each one of the other lower shaft gears coupled with the upper shaft gear, to obtain representative samples from a gradual fault progression. The difference between reference samples from normal and progressing fault conditions were employed to set simple threshold levels for each one of the vibration parameters to distinguish between different conditions. More advanced signal processing and pattern recognition techniques can be employed instead. However, the focus in this example is on offering a visual analytics view of the product, based on the data processing chain and not the exact signal and pattern analysis.

The visualisation application was developed in the *Processing* environment (processing.org), an Open Source Development Environment for Interactive Visualisation. The application presents a range of options for interactive visualisation. The application can produce reports and visual analytics graphs for the raw signal, the measured parameters and the FFT of the raw vibration signal, as well as motor temperature (Fig. 19).

The comparison of threshold values estimated from reference data and parameters extracted from subsequent observations is passed to the visualization layer of the application. This offers a 3D model of the test rig highlighting visual features by colours, conveying contextual meaning. For example, sensor locations are marked in blue colour. The diagnosis outcome is communicated by superimposing fault conditions features on the 3D CAD product representation, wherein mechanical components are highlighted in red to indicate faulty condition. Such visual features can be seen in an example screen captured from the visualisation application (Fig. 20).

Typical monitoring systems already convey measurement data and faults to users. However, blending visual features in 3D product representations offers an additional HIL option, further aiding a user to interact with product relevant data in a way relevant to non-monitoring contexts, such as when reviewing historical data and FMECA knowledge (Pistofidis et al., 2016). In this application case a user interacting with the application is able to access maintenance linked knowledge which is naturally more actionable, as it is shared in a contextually relevant way. As an example, a user handling FMECA knowledge is supported with visual features to understand the context of timelines of knowledge-rich events, and is thereby better aided to perform a FMECA revision cycle (Pistofidis et al., 2016). The use can switch faster between different product, knowledge, and measurement views, making the review of relevant history data, events, and FMECA knowledge more directly actionable. Thus, a design-stage tool, namely FMECA, is looked upon together with MoL data and disseminated in visually relevant ways, contributing to upper layer context management, ie context dissemination (Perera et al., 2014).

5. Conclusion and discussion

This paper makes the case for more efficient integration of human and non-human actors in sociotechnical systems. It argues that for the potential added value of bringing together human and non-human actors to be amplified, research and development solutions are required to allow not only a mere physical and technological integration, but to identify also ways of creating seamless data, information, knowledge, and decision flows. To address such needs, this paper introduced novel concepts, methods, and tools for enabling the engagement of human in the loop of sociotechnical systems. Seeking to make such HIL integration more effective, the adopted approach takes into account what makes a human actor unique, namely human cognitive capabilities, and sought to align developed methods and tools with such capabilities. Linked with such cognitive capabilities are some key design aspects of technology enablers introduced in this paper, including HIL for linked data and knowledge, HIL in the machine learning loop, and human in the loop of Visual Analytics. The research reported in this paper applied such concepts to application cases relevant to production environments, which had some common condition monitoring elements but different functional requirements and technology implementations.

The key aim of the application demonstrators was to show how some of the key underlying concepts for HIL integration in CPS are served by bringing together key Industry 4.0 technologies, such as Internet of Things, Visual Data Analytics and Machine Learning in asset and product lifecycle management application contexts. At the lower end of an IoT architecture, this paper introduced an example of an abstract IoT edge node design, which can support self-awareness of connected assets, while benefitting from HIL interaction in machine learning. In order to highlight the versatility of the approach, the abstraction architecture was instantiated with different hardware and software implementation options, covering low, mid, and high end requirements. At the level of human interaction with technical systems, the proposed concepts were mapped into designs and application demonstrators for enabling cloud-oriented implementations of metadata management and visual analytics, bringing together data, annotations, and knowledge relevant to the roles of such actors in production environments and in asset and product lifecycle management.

The presented research contributes to a growing body of literature on sociotechnical industrial environments. Further research in a number of directions to unlock more intelligent capabilities, as well as to increase adoption and improve performance in industrial sociotechnical environments is needed. The usage of knowledge graphs needs to be further formalised and supported by methods and tools to manage the evolution and reorganization of their connections and the learning of their association strengths. Appropriate semantic modelling needs to target both broad and application specific contexts and drive context reasoning. Furthermore, context modelling and reasoning need appropriate methods to manage evolving contexts, based on learning from actual human and non-human actors interaction, so as to meet not only the typical challenges in big data environments, but also feature additional capabilities for handling also human-contributed micro-knowledge. While machine learning approaches can be very effective in relatively well defined problems, or in cases where there is a wealth of data, production environments often suffer in both aspects and effective methods for HIL in the machine learning loop for such real application problems need further research attention. The effective integration of human cognitive capabilities in technical environments is still a long way off from delivering sociotechnical system with integrated cognitive capabilities and further studies need to focus on such aspects of technology enabled future cognitive factories.

Ultimately, successful integration of human and technical systems depends not only on unlocking new capabilities and improving performance, but also on the acceptance and adoption of the offered solutions in sociotechnical systems. This produces a need for introducing performance criteria and metrics, as well as conducting research for understanding the underlying factors that can contribute to improved success and adoption of Industry 4.0 sociotechnical systems. While technology enablers will be gradually maturing, non-technological barriers would still need to be addressed. This paper focused on the development of methods and tools that address functional requirements in such environments. However, non-functional features, related to security, ethics, and privacy, deserve much attention, in order to make any such solutions acceptable to industry.

Acknowledgments

The research was partially supported by the Cranfield University Innovate UK Design Rule KTP (PR 510841) grant “IoT and Big Data in PLM” and the ATHENA Research Centre research grants WelCOM - Wireless Sensor Networks for Optimal Asset Lifecycle Management (GSRT 09SYN-71-856) and PALMIS - Product and Asset Lifecycle Management Infrastructure and Services (GSRT/IKY/SIEMENS 2014–2016). The contribution towards the experimental part of the work of Cranfield University researchers Stefano Tedeschi and Maurizio Bevilacqua (in the Visual Analytics demonstrator), as well as of Mr Angelos Papadopoulos from Kleemann Lifts on the lifts demonstrator, is acknowledged.

References

Alexakos, C., Anagnostopoulos, C., Fournaris, A., Kalogeras, A., & Koulamas, C. (2017). Production process adaptation to IoT triggered manufacturing resource failure events. In *IEEE international conference on emerging technologies and factory automation, ETFA* (pp. 1–8). https://doi.org/10.1109/ETFA.2017.8247671.

Arica, E., Oliveira, M., & Emmanouilidis, C. (2018). ‘Performance Measurement in Sensorized Sociotechnical Manufacturing Environments. In I. Moon, G. M. Lee, J. Park, D. Kiritsis, & G. von Cieminski (Eds.), *Advances in Production Management Systems. Smart Manufacturing for Industry 4.0* (pp. 172–180). Springer.

Cassina, J., Taisch, M., Potter, D., & Parlikad, A. K. (2008). Development of PROMISE architecture and PDKM semantic object model. In *2008 IEEE International Technology Management Conference (ICE), (i)* (pp. 1–8).

Colombo, A. W., Karnouskos, S., Shi, Y., Yin, S., & Kaynak, O. (2016). Industrial cyber-physical systems. *Proceedings of the IEEE, 104*(5), 899–903. https://doi.org/10.1109/JPROC.2016.2548318.

da Silva, M. J., Pereira, C. E., & Gotz, M. (2016). Knowledge awareness in alarm system database. In *3rd IFAC workshop on advanced maintenance engineering, services, and technology, AMEST 2016, Biarritz, France, IFAC Papers Online: 49(28)* (pp. 297–302).

El Kadiri, S., Grabot, B., Thoben, K. D., Hribernik, K., Emmanouilidis, C., et al. (2016). Current trends on ICT technologies for enterprise information systems. *Computers in Industry, 79*, 14–33. https://doi.org/10.1016/j.compind.2015.06.008.

Emmanouilidis, C., Beroncelj, L., Bevilacqua, M., Tedeschi, S., & Ruiz-Carcel, C. (2018). Internet of things – enabled visual analytics for linked maintenance and product lifecycle management. In *INCOM 2018, 16th IFAC symposium on information control problems in manufacturing: 51* (pp. 435–440). 11-13 June 2018, IFAC Papers Online.

Emmanouilidis, C., Pistofidis, P., Fournaris, A., Bevilacqua, M., Durazo-Cardenas, I., Botsaris, P. N., et al. (2016). Context-based and human-centred information fusion in diagnostics. In *3rd IFAC workshop on advanced maintenance engineering, services, and technology, AMEST 2016: 49* (pp. 220–225). *IFAC-PapersOnLine*.https://doi.org/10.1016/j.ifacol.2016.11.038.

Endert, A., Hossain, M. S., Ramakrishnan, N., North, C., Fiaux, P., & Andrews, C. (2014). The human is the loop: new directions for visual analytics. *Journal of Intelligent Information Systems, 43*(3), 411–435. https://doi.org/10.1007/s10844-014-0304-9.

Fantini, P., Pinzone, M., & Taisch, M. (2018). Placing the operator at the centre of Industry 4.0 design: Modelling and assessing human activities within cyber-physical systems. *Computers and Industrial Engineering*. In Press https://doi.org/10.1016/j.cie.2018.01.025.

Fantini, P., Tavola, G., Taisch, M., Barbosa, J., Leitao, P., Liu, Y., et al. (2016). Exploring the integration of the human as a flexibility factor in CPS enabled manufacturing environments: Methodology and results. In *IECON proceedings (industrial electronics conference)* (pp. 5711–5716). https://doi.org/10.1109/IECON.2016.7793579.

Feng, S., Setoodeh, P., & Haykin, S. (2017). Smart home: cognitive interactive people-centric internet of things. *IEEE Communications Magazine, 55*(2), 34–39. https://doi.org/10.1109/MCOM.2017.1600682CM.

Framling, K., Kubler, S., & Buda, A. (2014). Universal messaging standards for the IoT from a lifecycle management perspective. *IEEE Internet of Things Journal, 1*(4), 319–327. https://doi.org/10.1109/JIOT.2014.2332005.

Gross, K. C., Baclawski, K., Chan, E. S., Gawlick, D., Ghoneimy, A., & Liu, Z. H. (2017). A supervisory control loop with prognostics for human-in-the-loop decision support and control applications. In *2017 IEEE conference on cognitive and computational aspects of situation management (CogSIMA)* (pp. 27–31). March.

Idreos, S., Papaemmanouil, O., & Chaudhuri, S. (2015). Overview of Data Exploration Techniques. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data - SIGMOD '15* (pp. 277–281). https://doi.org/10.1145/2723372.2731084.

Jun, H. B., Kiritsis, D., & Xirouchakis, P. (2007). Research issues on closed-loop PLM. *Computers in Industry, 58*(8–9), 855–868. https://doi.org/10.1016/j.compind.2007.04.001.

Kaplan, R. S., Norton, D., Leadership, Z., Mayo, A. J., Nohria, N., Allmendinger, G., et al. (2005). Four strategies for smart services. *Harvard Business Review, 83*(10), 131–145. https://doi.org/10.1225/R0510J.

Katsouros, V., Koulamas, C., Fournaris, A. P., & Emmanouilidis, C. (2015). Embedded event detection for self-aware and safer assets. IFAC symposium on fault detection, supervision and safety for technical processes SAFEPROCESS 2015, Paris, 2–4 September. *IFAC-PapersOnLine, 28*(21), 802–807. https://doi.org/10.1016/j.ifacol.2015.09.625.

Kiritsis, D. (2011). Closed-loop PLM for intelligent products in the era of the Internet of things. *CAD Computer Aided Design, 43*(5), 479–501. https://doi.org/10.1016/j.cad.2010.03.002.

Kiritsis, D., Jun, H.-B., & Xirouchakis, P. (2008). Closing product information loops with product-embedded information devices: RFID technology and applications, models and metrics. In S. B. Miles, S. E. Sarma, & J. R. Williams (Eds.), *RFID Technology and Applications* (pp. 169–182). Cambridge: Cambridge University Press.

Kiritsis, D., Bufardi, A., & Xirouchakis, P. (2003). Research issues on product lifecycle management and information tracking using smart embedded systems. *Advanced Engineering Informatics, 17*(3–4), 189–202. https://doi.org/10.1016/j.aei.2004.09.005.

Kubler, S., Yoo, M.-J., Cassagnes, C., Framling, K., Kiritsis, D., & Skilton, M. (2015). Opportunity to leverage information-as-an-asset in the IoT – The road ahead. In 2015 3rd international conference on future internet of things and cloud, *(september 2014* (pp. 64–71). https://doi.org/10.1109/FiCloud.2015.63.

Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. *Cognitive Systems Research, 10*(2), 141–160. https://doi.org/10.1016/j.cogsys.2006.07.004.

Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & Siegel, D. (2014). Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. *Mechanical Systems and Signal Processing, 42*(1–2), 314–334. https://doi.org/10.1016/j.ymssp.2013.06.004.

Leitão, P., Karnouskos, S., Ribeiro, L., Lee, J., Strasser, T., & Colombo, A. ,W. (2016). Smart agents in industrial cyber-physical systems. *Proceedings of the IEEE, 104*(5), 1086–1101. https://doi.org/10.1109/JPROC.2016.2521931.

Leitão, P., Member, S., Ma, V., & Vrba, P. (2013). Past, present, and future of industrial agent applications. *IEEE Transactions on Industrial Informatics, 9*(4), 2360–2372. https://doi.org/10.1109/TII.2012.2222034.

Liyanage, J. P., Lee, J., Emmanouilidis, C., & Ni, J. (2009). Integrated E-maintenance and intelligent maintenance systems. In *Handbook of maintenance management and engineering* (pp. 499–544). https://doi.org/10.1007/978-1-84882-472-0_20.

Ma, M., Lin, W., Pan, D., Wang, P., Zhou, Y., & Liang, X. (2017). Data and decision intelligence for internet of things: Putting human in the loop. In *2017 IEEE 3rd international conference on big data security on cloud* (pp. 190–195). IEEE. https://doi.org/10.1109/BigDataSecurity.2017.49.

Matsokis, A., & Kiritsis, D. (2010). Computers in Industry An ontology-based approach for Product Lifecycle Management. *Computers in Industry, 61*(8), 787–797. https://doi.org/10.1016/j.compind.2010.05.007.

McFarlane, D., Sarma, S., Chirn, J. L., Wong, C. Y., & Ashton, K. (2002). The intelligent product in manufacturing control and management. *IFAC Proceedings Volumes, 35*(1), 49–54. https://doi.org/10.3182/20020721-6-ES-1901.00011.

Meyer, G. G., Främling, K., & Holmström, J. (2009). Intelligent Products: A survey. *Computers in Industry, 60*(3), 137–148. https://doi.org/10.1016/j.compind.2008.12.005.

Nunes, D. S., Zhang, P., & Silva, J. S. (2015). A survey on human-in-the-loop applications towards an internet of all. *IEEE Communications Surveys & Tutorials, 17*(2), 944–965.

Patterson, R. E. (2017). Intuitive cognition and models of human-automation interaction. *Human Factors, 59*(1), 101–115. https://doi.org/10.1177/0018720816659796.

Perera, C., Zaslavsky, A., Christen, P., & Georgakopoulos, D. (2014). Context aware computing for the internet of things: A survey. *IEEE Communications Surveys and Tutorials, 16*(1), 414–454. https://doi.org/10.1109/SURV.2013.042313.00197.

Pistofidis, P., Emmanouilidis, C., Papadopoulos, A., & Botsaris, P. N. (2016). Management of linked knowledge in industrial maintenance. *Industrial Management & Data Systems, 116*(8), 1741–1758. https://doi.org/10.1108/IMDS-10-2015-0409.

Pistofidis, P., Emmanouilidis, C., Koulamas, C., Karampatzakis, D., & Papathanassiou, N. (2012). A layered e-maintenance architecture powered by smart wireless monitoring components. In *2012 IEEE International Conference on Industrial Technology, ICIT 2012* (pp. 390–395). https://doi.org/10.1109/ICIT.2012.6209969.

Romero, D., Stahre, J., Wuest, T., Noran, O., Bernus, P., Fast-berglund, Å., et al. (2016). Towards an operator 4.0 typology: A human-centric perspective on the fourth industrial revolution technologies. *Computers and indistrial engineering proceedings, CIE46*.

Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust regression and outlier detection*. New York: Wiley & Sons Inc.

Snidaro, L., García, J., & Llinas, J. (2015). Context-based information fusion: A survey and discussion. *Information Fusion, 25*, 16–31. https://doi.org/10.1016/j.inffus.2015.01.002.

Subramania, H. S., & Khare, V. R. (2011). Pattern classification driven enhancements for human-in-the-loop decision support systems. *Decision Support Systems, 50*(2), 460–468. https://doi.org/10.1016/j.dss.2010.11.003.

Thomas, J. J., & Cook, K. A. (2005). Illuminating the path: The research and development agenda for visual analytics. *IEEE Computer Society*, 184. https://doi.org/10.3389/fmicb.2011.00006.

Tuckey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics, 33*(1), 1–67.

Tuckey, J. W. (1977). Exploratory data analysis. *Reading*. Mass: Addison-Wesley.

Vathy-Fogarassy, Á., & Hugyák, T. (2017). Uniform data access platform for SQL and NoSQL database systems. *Information Systems, 69*, 93–105. https://doi.org/10.1016/j.is.2017.04.002.

Vermesan, O., & Friess, P. (Eds.) (2016). *Digitising the industry: internet of things connecting the physical, digital and virtual worlds*. River Publishers.

Vermesan, O., Friess, P., Guillemin, P., Gusmeroli, S., Sundmaeker, H., Bassi, A., et al. (2011). Internet of things strategic research roadmap. In *Internet of things - global technological and societal trends from smart environments and spaces to green ICT* (pp. 9–52). River Publishers.

Willig, A. (2008). Recent and emerging topics in wireless industrial communications: A selection. *IEEE Transactions on Industrial Informatics, 4*(2), 102–122. https://doi.org/10.1109/TII.2008.923194.

Wuest, T., Hribernik, K., & Thoben, K. D. (2015). Accessing servitisation potential of PLM data by applying the product avatar concept. *Production Planning and Control, 26*(14–15), 1198–1218. https://doi.org/10.1080/09537287.2015.1033494.

Yoo, M., Grozel, C., & Kiritsis, D. (2016). Closed-loop lifecycle management of service and product in the internet of things: semantic framework for knowledge integration. *Sensors, 17*(7). https://doi.org/10.3390/s16071053.

2019-03-19

Enabling the human in the loop: Linked data and knowledge in industrial cyber-physical systems

Emmanouilidis, Christos

Elsevier

Emmanouilidis C, Pistofidis P, Bertoncelj L, et al., (2019) Enabling the human in the loop: Linked data and knowledge in industrial cyber-physical systems. Annual Reviews in Control, Volume 47, 2019, pp. 249-265

https://doi.org/10.1016/j.arcontrol.2019.03.004